

OM of: US-08-962-560A-4 to: EST: * out_format : pfs
Date: Feb 4, 2000 8:49 PM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+g2n.model -DEV=xip
-O=/cgm1_1/USPO.spool/US08962560/runat_03022000_141841_3589/app-query.fasta.1
-DB=FAST -OEMT=fastcap -SUFFIX=first -GAPOP=12.000 -GAPEXT=4.500
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
-DELXT=7.000 -START=1 -MATRIX=Diosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=PCT -ALIGN=15 -MODE=LOCAL
-OUTMT=pfs -NORM=ext -MINLEN=0 -MAXLEN=1000000 -USER=US08962560
-NCPU=6 -ICPU=3 -NO_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-08-962-560A-4
Query length: 212
Database: EST: *
Database sequences: 4538634
Database length: 1887831982
Search time (sec): 15337.100000

score_list:	Strid	Orig	ZScore	Escore	Len	Documentation
gb_est20:AA849015	-	843.00	1186.00	2.7e-57	623	AA849015 EST191777 Normalized
gb_est38:AM025552	-	819.50	1155.02	1.4e-55	500	AM025552 wu57904.x1 NC1 CGAP
gb_est27:AI140024	-	811.50	1142.17	7.1e-55	591	AI140024 lg88d01.x1 Soares_NLM
gb_est38:AM082710	-	810.50	1142.51	7.1e-55	492	AM082710 xb61f03.x1 Soares_NLM
gb_est31:AI683421	-	808.50	1140.01	9.8e-55	476	AI683421 tw51f03.x1 NC1 CGAP
gb_est34:AI806657	-	767.50	1081.80	1.7e-51	502	AI806657 wf35f03.x1 Soares_NLM
gb_est20:AA881502	-	759.00	1068.11	1.0e-50	602	AA881502 vx19f05.x1 Soares_NLM
gb_est27:AA957508	-	757.00	1068.59	9.4e-51	426	AA957508 UT-R-EL-gb-f-05-0-UT-R
gb_est38:AM047382	-	754.00	1064.45	1.6e-50	422	AM047382 UT-M-BH1-eme-e-06-0-UT
gb_est36:AI873269	-	747.00	1053.70	6.3e-50	464	AI873269 wf41c08.x1 Soares_NLM
gb_est11:AA212316	-	722.00	1016.81	7.2e-48	555	AA212316 mu35f12.x1 Soares_NLM
gb_est28:AI505924	-	718.00	1013.90	1.0e-47	417	AI505924 vx69g08.x1 Knowles Sol
gb_est38:AM048541	-	687.00	970.93	2.6e-45	389	AM048541 UT-M-BH1-ax-c-08-0-UT
gb_est28:AI502911	-	680.00	960.98	9.1e-45	393	AI502911 UT-R-CL-Kn-f-07-0-UT-R
gb_est16:AA57459	-	607.00	856.43	6.1e-39	475	AA57459 vx69g08.x1 Knowles Sol
gb_est32:AI710454	-	600.00	849.52	1.5e-38	349	AI710454 UT-R-A61-sak-b-01-0-UT
gb_est22:AI056051	-	592.50	835.31	9.2e-38	512	AI056051 ox46c05.x1 Soares_NLM
gb_est39:AM19108	-	582.00	821.05	5.7e-37	485	AM19108 mp66d03.x1 Soares_NLM
gb_est38:AM051817	-	576.50	812.41	1.7e-36	533	AM051817 w203h05.x1 NC1 CGAP
gb_est38:AM055217	-	568.50	800.82	7.7e-36	552	AM055217 w202e02.x1 NC1 CGAP
gb_est20:AA828319	-	558.50	790.35	7.2e-35	378	AA828319 ob55f04.x1 NC1 CGAP
gb_est37:AI968472	-	543.00	770.36	3.8e-34	312	AI968472 wt89e01.x1 NC1 CGAP
gb_est37:AI990702	-	543.00	770.36	3.8e-34	312	AI990702 ws31f09.x1 NC1 CGAP
gb_est31:AI671628	-	543.00	770.36	3.8e-34	312	AI671628 wt89e01.x1 NC1 CGAP
gb_est39:AM136461	-	521.00	739.40	2.0e-32	312	AM136461 UT-R-BH1-gd1-h-09-0-UT
gb_est24:AI215457	-	514.00	729.46	7.3e-32	315	AI215457 qn11d04.x1 Soares_NLM
gb_est24:AI243555	-	497.00	706.52	1.4e-30	284	AI243555 qh2e02.x1 Soares_NLM
gb_est18:AA020019	-	466.50	658.34	8.8e-28	493	AA020019 mh48c11.x1 Soares_NLM
gb_est30:AI939622	-	448.00	638.21	6.8e-27	265	AI939622 vx59c03.x1 Soares_NLM
gb_est19:AA116407	-	446.50	631.36	2.1e-26	436	AA116407 mp99a03.x1 Soares_NLM
gb_est12:AA83034	-	442.00	628.16	3.0e-26	299	AA83034 z16c101.x1 NC1 CGAP
gb_est10:AA145823	-	435.50	615.14	1.7e-25	471	AA145823 mq88g10.x1 Soares_NLM
gb_est31:AI701627	-	432.00	615.14	1.7e-25	281	AI701627 tz98c02.x1 NC1 CGAP
gb_est20:AA821494	-	422.00	597.01	1.7e-24	431	AA821494 vx59c03.x1 Soares_NLM
gb_est32:AI742049	-	418.00	588.86	4.9e-24	560	AI742049 wq38d03.x1 Soares_NLM
gb_est27:AI454298	-	408.00	582.92	1.0e-23	237	AI454298 UT-R-BO-ce-e-11-0-UT-R
gb_est19:AA118286	-	363.00	515.72	5.8e-20	358	AA118286 mp88g10.x1 Soares_NLM
gb_est20:AA863619	-	363.00	480.42	5.4e-18	360	AA863619 vx06a09.x1 Soares_NLM
gb_est20:AA884657	-	298.00	425.42	6.3e-15	316	AA884657 vx06a09.x1 Soares_NLM
gb_est15:AA485510	-	261.00	377.12	2.8e-12	198	AA485510 vx30a12.x1 Soares_NLM
gb_est19:AA789818	-	258.00	368.15	9.7e-12	350	AA789818 vx30a12.x1 Soares_NLM
gb_est38:AM019268	-	254.50	355.89	4.7e-11	756	AM019268 fd51a04.x1 Zebrafish W

gb_est39:AM141012 + 248.00 356.24 1.3e-11 279 | AM141012 EST291035 Normalized
gb_est26:AI339592 + 233.00 330.22 1.3e-09 467 | AI339592 qk67g11.x1 NC1 CGAP
gb_est19:AA756393 + 231.00 330.29 1.2e-09 345 | AA756393 vv42f06.x1 Soares 2

seq_name: gb_est20:AA849015

seq_documentation block:

LOCUS AA849015 623 bp mRNA EST 30-APR-1998
DEFINITION EST191777 Normalized rat lung, Bento Soares Rattus sp. cDNA clone
R1DJM34.3' end, mRNA sequence.
ACCESSION AA849015
VERSION AA849015.1 GI:2936555
KEYWORDS EST
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 623)
Lee N.H., Glodex A., Chandra, I., Mason, T.M., Quackenbush, J.,
Kerlavage, A.R. and Adams, M.D.
Rat Genome Project: Generation of a Rat EST (REEST) Catalog & Rat
Gene Index

JOURNAL Unpublished (1998)
COMMENT On Jan 9, 1998 this sequence version replaced gi:936155.
Contact: Lee, NH
ATCC

FEATURES
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13-21.

FEATURES

source location/qualifiers
1..623
/organism="Rattus sp."
/db_xref="ATCC (Inhost):2008043"
/db_xref="taxon:10118"
/clone="R1DJM34"
/clone_lib="Normalized rat lung, Bento Soares"
/note="Organ: Lung; Vector: pT73Pac; Site:1; EcorI;
Site:2; NotI"

BASE COUNT 94 a 193 c 254 g 82 t

alignment_scores:

Quality: 843.00 Length: 163
Ratio: 5.204 Gaps: 0
Percent Similarity: 99.387 Percent Identity: 97.546

alignment_block:
US-08-962-560A-4 x AA849015/rev ..

Align seg 1/1 to reverse of: AA849015 from: 1 to: 623

1 MetValAlaArgSngInValAlaAlaAspAsnAlaIleSerProAlaAl 17
|||||
491 ATGTGATGACAGTATACAGGTGGACGACCAATGATCATCTCCCGGCATC
17 agtuprorgargargsergluprroserSerSerSerSerSer 34
|||||
441 AGAGCCCGGACGCGCGCCAGACCATCTCTCTCTCTCTCTCTCTCGC 392
|||||
34 roAlaAlaProValArgProArgProArgProArgProAlaAlaProAla 50
|||||
391 CGCT 342
|||||
51 ProGlaAspPhrIstPhaArgThrPhaArgSerHisSerAspPyrrArg 67
|||||
341 CGCGGCGACACCTACCTCCGACCTCCGACCTCCGACCTCCGACCTCCG 292
|||||
67 giletrargThrSerAlaIleuAspAlaCysGlyPheTyrrGlyP 84
|||||

291 CATCAGCGGACGACGCTCTCTGAGCGCTTCTACTGGGGAC 242

84 roLeuSerValHisGlyAlaHisGlyLeuArgAlaGluProValGly 100
|||||
241 CCTGTGAGCGGATGGGGCCGACGAACGGGTGGTCCGAGCCCGTGGG 192
101 ThrPheLeuValArgAspSerArgGlnArgAsnCysPhePheAlaLeu 117
|||||
191 ACCTTCTGTGGTGGCGACAGTCCGACGGAACTGCTCTTCGCGCTCAG 142

117 rValysMetAlaSerGlyProThrSerIleArgValHisPheGlnAlaG 134
|||||
141 CGTGAAGATGGCTTCGGGCCCCGACGACATTCGTGCACTTCACAGCCG 92

134 lYArgPheHisIleuAspGlySerArgGluThrPheAspCysLeuPheGlu 150
|||||
91 GCGGCTTCCACCTGGAGCGGACGCGGAGACCTTCGACTGCTTCGAG 42

151 LeuLeuGluHisTyrValAlaAlaProArgArgMetLeu 163
|||||
41 CTGCTGGAGCACTACGTGGCGGCGCCCGCCGCACTGTG 3

seq_name: gb_est38:AM025552

seq_documentation_block: 500 bp mRNA 27-OCT-1999
LOCUS AM025552
DEFINITION W07904.X1 NCI-CGAP_K143 Homo sapiens cDNA clone IMAGE:2528022 3'
similar to TR:035716 035716 CYTOKINE INDUCIBLE SH2-CONTAINING
PROTEIN 7 ; mRNA sequence.
ACCESSION AM025552
VERSION AM025552.1 GI:5879082
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 500)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jun 5, 1998 this sequence version replaced gi:3187655.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LMW at:
www-bio.liml.gov/dbtrp/image/image.html

FEATURES
SOURCE
Seq primer: -40UP from G1bco
High quality sequence stop: 464.
Location/Qualifiers
1..500
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2528022"
/clone_1lb="NCI CGAP_K143"
/lab_host="DH10B"
/note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer,
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. mRNA
source: 2 pooled kidneys. Library went through one round
of normalization. Library constructed by Bento Soares and
M. Fatima Bonaldo."

BASE COUNT 76 a 157 c 198 g 68 t 1 others

ORIGIN

alignment_scores:
Quality: 819.50 Length: 165
Ratio: 5.059 Gaps: 1
Percent Similarity: 98.182 Percent Identity: 94.545

alignment_block:
US-08-962-560a-4 x AM025552/rev ..

Align seg 1/1 to reverse of: AM025552 from: 1 to: 500

1 MetValAlaArgAsnGlnValAlaAlaAspAsnAlaIleSerProAla 17
|||||
492 ATGTGTCACACACACACAGGTGGACCGACGACATTCGATTCACACAGCACC 443

17 acGluProArgArgArgSerGluProSerSerSerSerSerSerSer 34
|||||
442 AGAGCCCCGAGCGGCGCCAGAACN...TCCCTCTTCTCTCTCTCTGCG 396

34 roAlaAlaProValArgProArgProArgProAlaValProAlaProAla 50
|||||
395 CCGCGGCCCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 346

51 ProGlyAspThrHisPheArgThrPheArgSerHisSerAspTyrArg 67
|||||
345 CCGGCGACACGCACTTCGACATTCGATTCGACGCGCGGATTCAGCGGG 296

67 glierThrArgThrSerAlaLeuLeuAspAlaCysGlyPheTyrTrpGly 84
|||||
295 CATCAGCGCGCGCAGCGGCGCTCTGAGCGCTCGCGGATTCAGCGGGG 246

84 roLeuSerValHisGlyAlaHisGlyLeuArgAlaGluProValGly 100
|||||
245 CCTGTGAGCTGCACGCGGGGCGACGAGCGGTGCGCGCGAGCCCTGGGC 196

101 ThrPheLeuValArgAspSerArgGlnArgAsnCysPhePheAlaLeu 117
|||||
195 ACCTTCTGTGGTGGCGACAGCCCGCAGCGGAATCTTTTGGCCCTTAG 146

117 rValysMetAlaSerGlyProThrSerIleArgValHisPheGlnAlaG 134
|||||
145 CGTGAAGATGGCTTCGGGACCCAGACGATCCGCGCTTCAGCGCGG 96

134 lYArgPheHisIleuAspGlySerArgGluThrPheAspCysLeuPheGlu 150
|||||
95 GCGGCTTTCACCTGGATGGACGCCGAGAGCTTGACTGCTCTTCGAG 46

151 LeuLeuGluHisTyrValAlaAlaProArgArgMetLeuGlyAla 165
|||||
45 CTGCTGGAGCACTACGTGGGCGCGCGCGCGCGCACTGCGGGGCC 1

seq_name: gb_est27:A1400249

seq_documentation_block: 591 bp mRNA 30-MAR-1999
LOCUS A1400249
DEFINITION t968001.x1 Soares_NhmPpu.S1 Homo sapiens cDNA clone IMAGE:2113921
3 similar to TR:015097 015097 JAB, PATITIL CDS ;contains MER22.B2
MSR1 repetitive element ; mRNA sequence.
ACCESSION A1400249
VERSION A1400249.1 GI:4243336
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 591)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Mar 20, 1998 this sequence version replaced gi:2980432.


```

1 MetValAlaArgAsnGlnValAlaAlaAspAsnAlaIleSerProAlaAl 17
492 ATGTACACACACACAGGTGGCGACGACATCACTCTCCACAGACG 443
17 AGUProArgArgArgSerGluProSerSerSerSerSerSerSer 34
442 AGAGCCCCGAGCGGGCCAGACCT...TCTCTCTCTCTCTCTCTCTCCGCG 396
34 rolaAlaProValArgProArgProCysProAlaValProAlaProAla 50
395 CCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 346
51 ProGlyAspThrHisPheArgThrPheArgSerHisSerAspTyrArg 67
345 CCGGGCACAGCAGCTTCGACATTCGACGCGCCGACCGCATTCGCGGCG 296
67 gileThArgThrSerAlaLeuLeuAspAlaCysGlyPheTyrTrpGly 84
295 CATCAGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 246
84 roLeuSerValHisGlyAlaHisGlyAlaArgLeuArgAlaGluProValGly 100
245 CCGTACGCGTGCACGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 196
101 ThrPheLeuValArgAspSerArgGlnArgAsnGlyPhePheAlaLeu 117
195 ACCTTCCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 146
117 rValIleMetAlaSerGlyProThrSerIleArgValHisPheGlnAla 134
145 CGTGAAGATGGCTCGGGGACCCAGACATCCGGTGCACCTTGAAGCGG 96
134 lYArgPheHisLeuAspGlySerArgGluThrPheAspCysLeuPheGlu 150
95 GCCGCTTACCTGATGGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 46
151 leuLeuGlnHisTyrValAlaAlaProArgArgMetLeuGlyAla 165
45 CTGTGGAGCAGCTACGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1
seq_name: gb_est31:AI683421

```

```

seq_documentation_block: 476 bp mRNA EST 26-MAY-1999
LOCUS AI683421 tw51a09.x1 NCI-CGAP-Utl1 Homo sapiens cDNA clone IMAGE:2263192 3'
DEFINITION Similar to TR:015097 015097 JAB, PARTIAL CDS ;, mRNA sequence.
ACCESSION AI683421 GI:4893603
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 476)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On May 18, 1998 this sequence version replaced gi:3137658.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Koskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbfp/image/image.html

```

Seq primer: -40UP from G1bco

High quality sequence stop: 433.

FEATURES

Location/Qualifiers

1..476
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2263192"
/clone_11b="NCI-CGAP-Utl1"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-Sport6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Life Technologies catalog #:
11538-014"

BASE COUNT 70 a 154 c 197 g 54 t 1 others

alignment_scores:

Quality: 808.50 Length: 159
Ratio: 5.150 Gaps: 1
Percent Similarity: 98.742 Percent Identity: 96.226

alignment_block:

US-08-962-560A-4 x AI683421/rev ..

Align seg 1/1 to reverse of: AI683421 from: 1 to: 476

```

16 AlAlaGluProArgArgSerGluProSerSerSerSerSer 32
476 GCAGCAGAGCCCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 430
32 rSerProAlaAlaProValArgProArgProCysProAlaValProAla 49
429 CTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 380
49 rolaProGlyAspThrHisPheArgThrPheArgSerHisSerAspTyr 65
379 CGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 330
66 ArgArgIleThrArgThrSerAlaLeuLeuAspAlaCysGlyPheTyr 82
329 CGCGCATCACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 280
82 pGlyProLeuSerValHisGlyAlaHisGlyAlaArgLeuArgAlaGluPro 99
279 GGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 230
99 aGlyThrPheLeuValArgAspSerArgGlnArgAsnGlyPhePheAla 115
229 TGGGACCTTCCTGTGGCGGACAGCCGCGGCGGCGGCGGCGGCGGCGGCGG 180
116 leuSerValIleMetAlaSerGlyProThrSerIleArgValHisPheG 132
179 CTGAGCGTGAAGATGGCTCGGGACCCAGACATCCGGTGCACCTTGA 130
132 naAGlyArgPheHisLeuAspGlySerArgGluThrPheAspCysLeu 149
129 GGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 80
149 heGluLeuGlnHisTyrValAlaAlaProArgArgMetLeuGlyAla 165
79 TCAGAGCGTGCAGCAGCTACGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 30
166 ProLeuArgGlnArgArgValArgPro 174
29 CCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3
seq_name: gb_est34:AI806697

```

```

seq_documentation_block: 502 bp mRNA EST 07-JUL-1999
LOCUS AI806697 tw35f03.x1 Soares_NFL_T_GBC_s1 Homo sapiens cDNA clone
DEFINITION

```

```

ACCESSION      IMAGE:2357597 3' similar to TR:015097 015097 JAB, PARTIAL CDS ;
AIB06697
VERSION        AIB06697.1
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
                Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      NCI-CCAP http://www.nccl.nlm.nih.gov/nciccap.
AUTHORS        1 (bases 1 to 502)
TITLE          National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
                Tumor Gene Index
                Unpublished (1997)
JOURNAL        On Jun 5, 1998 this sequence version replaced gi:1187434.
COMMENT        Contact: Robert Strausberg, Ph.D.
                Tel: (301) 496-1550
                Email: Robert_Strausberg@nih.gov
                This clone is available royalty-free through LINT ; contact the
                IMAGE Consortium (info@image.lln.gov) for further information.
                Seq primer: -40UP from Gldco
                High quality sequence stop: 457.
FEATURES
SOURCE         location/Qualifiers
                1..502
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:2357597"
                /clone_11b="Soares_NFL_T_GBC_S1"
                /lab_host="DH10B"
                /note="Organ: pooled; Vector: pT7AD-Pac (Pharmacia) with
                a modified polylinker. Site.1: Not I. Site.2: Eco RI;
                Equal amounts of plasmid DNA from three normalized
                libraries (fetal lung NDHL19W, testis NHT, and B-cell
                NCI-CGAP-GCB1) were mixed, and ss circles were made in
                vitro. Following HAP purification, this DNA was used as
                tracer in a subtractive hybridization reaction. The driver
                was PCR-amplified cDNAs from pools of 5,000 clones made
                from the same 3 libraries. The pools consisted of
                I.M.A.G.E. clones 297480-302087, 682632-687239,
                726408-728711, and 729096-731399. Subtraction by Bento
                Soares and M. Fatima Bonaldo."
BASE COUNT     78 a 157 c 198 g 69 t
ORIGIN
alignment_scores:
    Quality: 767.50          Length: 161
    Ratio: 4.984            Gaps: 2
    Percent Similarity: 95.652    Percent Identity: 91.925
alignment_block:
US-08-962-560A-4 x AIB06697/rev ..
Align seg 1/1 to reverse of: AIB06697 from: 1 to: 502
5 AsnglnValAlaAlaaspAsnaAlaIleSerProAlaAlaGluProArgar 21
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
482 CACGAGTGGCAGCCGACATGCAATGCACTCCACAGCAGCAGAGCCCGGAC 43
21 gATgSerGluProSerSerSerSerSerSerSerSerProAlaAlaProy 38
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
432 GCGGCGAGAACT...TCCTCCCTTCCTTCCTTCCTTCGCGCGCCCGC 386
38 AlarProArgProCysProAlaValProAlaProAlaProGlyAspThr 54
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
386 GCGCGCCGCGCGCGGCCCCGCGGTCCGCGCCCCCGCGCGCGCGCAGC 336
55 HisPheArgThrPheArgSerHisSerAspTyrArgArgIleThrArgn 71
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
335 CACTTCGCAATTCGTTCCGACCGCGAGTACCGCGCGCATCACGCGGCG 286
71 rSerAlaLeuLeuAspAlaCysGlyPheTyrTrpGlyProLeuSerValH 88
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

285 CAGCGGGCTCCTGACGCGTCCGATTCTACTGAGGGGCCCTGAGCGTCG 236
98 1AG1YALH1SGLUARGLEUAARG1AAGLUPROVALG1YHRPHELUVAL 104
|||||
235 ACGGGCGGCACAGAGCGGTGCGCGCGACCCGCTGGCACCTTCGTGGTG 186
105 ArgAspSerArgIlnArgAnCySPhePeAlaLeuSerVallyMetal 121
|||||
185 CGCGACACCGCCGACGGAATGCTTTTCGCCCTTAGCGTGAAGATGCC 136
121 aSeRgIProThrSerIleArGVAlHisPheGlnAlaGlyArgPHeIsl 138
|||||
135 CTCGGGACCACAGACATCCGGGTGCATCTTCAGCGCGCGCGCTTCACG 86
138 euASpeLyseArGrGUthrPheASPCyLeuPheGluLeuGluHIS 154
|||||
85 TGATGCGCAGCCGCGAGAGCTTCGACTGCTCTTCAGAGTGTGAGACAC 36
155 TyrValAlaAlaProArGrMetLeuGly 164
|||||
35 TACGTGGCGAAAGCCGCGCGACCTGCTGGG 5
seq_name: gb_est20:AA881502

seq_documentation_block:
LOCUS AA881502 602 bp mRNA EST 26-MAR-1998
DEFINITION vx19f05.1 Soares 2NbMT Mus musculus cDNA clone IMAGE:1264929 5'
similar to TR:Q35716 Q35716 SUPPRESSOR OF CYTOKINE SIGNALLING-1,
[1] ; mRNA sequence.
ACCESSION AA881502
VERSION AA881502.1 GI:2990812
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murineae; Mus.
REFERENCE 1 (bases 1 to 602)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucada,T., Lacy,M., Le,M., Martin,D., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theisinger,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1398104.
CONTACT: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
TEL: 314 286 1800
FAX: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:667481
Possible reversed clone: similarity on wrong strand
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 459.
location/Qualifiers
1..602
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1264929"
/clone_id="Soares 2NbMT"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Vector: pUT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I oligo(dT) primer [5',
TGTTCACAATCTGATGAGGAGCGCGCTTTTTTTTTTTTTTTTTT

```

3'): double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 76 a 243 c 181 g 101 t 1 others
ORIGIN

alignment_scores:
Quality: 759.00 Length: 147
Ratio: 5.234 Gaps: 0
Percent Similarity: 98.639 Percent Identity: 97.279

alignment_block:
US-08-962-560A-4 x AA881502 ..

Align seg 1/1 to: AA881502 from: 1 to: 602

```

1 MetValAlaArgAsnGlnValAlaAlaAspAsnAlaIleSerProAlaAl 17
|||||
162 ATGTACACAGCAACAGGTGACAGCGACATGCGATCCCGGAGC 211
|||||
17 AGUProArgArgSerGluProSerSerSerSerSerSerSer 34
|||||
212 AGAGCCCGGCGGCGGTACAGCCCTCTGCTCTCTGCTCTCCGCGC 261
|||||
34 roAlaAlaProValArgProArgProCysProAlaValProAlaProAla 50
|||||
262 CACGGGCGCCGCGTGGCGTCCCGGCGGCGGCGGCGGCGGCGGCGGCGG 311
|||||
51 ProGlyAspThrHisPheArgThrPheArgSerHisSerAspTyrArg 67
|||||
312 CCGTGGGAMCCTCCTCCGACCTTCCGCTCCGACATTCACATTCACGCG 361
|||||
67 gileThArgThrSerAlaLeuAspAlaCysGlyPheTyrTrpGlyP 84
|||||
362 CACACCGGACCAAGCGCGCTCTGAGCGGCTTATGGGAGC 411
|||||
84 roLeuSerValHisGlyAlaHisGlyAlaArgLeuArgAlaGluProValGly 100
|||||
412 CCTGAGCGGTGACAGCGGCGGCGGCGGCGGCTGCGTCCGAGCGGCG 461
|||||
101 ThrPheLeuValArgAspSerArgGlnArgAsnCysPheAlaLeu 117
|||||
462 ACCTTCTGTGGTGGCGACGTGCGACAGGAGCTTCTTCGGGCTCAG 511
|||||
117 ValIysMetAlaSerGlyProThrSerIleArgValHisPheGlnAlaG 134
|||||
512 GGTGAGATGCTTCGGGGCCCAAGACATCCGCTGCACTTCAGGCGC 561
|||||
134 LysArgPheHisLeuAspGlySerArgGluThrPheAspCys 147
|||||
562 GCCGCTTCCACTTGAGCGGCGGCGGCGGCTTCGACTGC 602
|||||

```

seq_name: gb_est27.AA957508

seq_documentation_block:

LOCUS AA957508 426 bp mRNA EST 09-MAR-1999
DEFINITION UI-R-EI-gb-f-05-0-UI.s1 UI-R-EI Rattus norvegicus cDNA clone
UI-R-EI-gb-f-05-0-UI 3', mRNA sequence.
ACCESSION AA957508
VERSION AA957508.1 GI:4277398
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 426)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704447
COMMENT On May 7, 1998 this sequence version replaced gi:3121203.

Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu

Oligo-dt track not found. Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.
Fatima Bonaldo, Ph.D. Clone distribution: clones will be available
through Research Genetics The following repetitive elements were
found in this cDNA sequence: 366-425, >GC-rich#low_complexity
Seq primer: M13 forward.

FEATURES

source

1..426

Location/Qualifiers

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-EI-gb-f-05-0-UI"

/clone_id="UI-R-EI"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; The UI-R-EI
library is a subtracted library derived from the UI-R-EI
library. The UI-R-EI library consisted of a mixture of
individually tagged normalized libraries constructed from
8, 12 and 18-day embryo. The tag is a string of 3-5
nucleotides present between the Not I site and the
oligo-dt track which allows identification of the library
of origin of a clone within the mixture. The subtracted
library (UI-R-EI) was constructed as follows: PCR
amplified cDNA inserts from a pool of UI-R-EI clones from
which 3' ESTs had been derived was used as a driver in a
hybridization with the UI-R-EI library in the form of
single-stranded circles. The remaining single-stranded
circles (subtracted library) was purified by
hydroxyapatite column chromatography, converted to
double-stranded circles and electroporated into DH10B
bacteria (Life Technologies) to generate the UI-R-EI
library. This procedure has been previously described
(Bonaldo, Lennon and Soares, Genome Research 6: 791-806,
1996)."

BASE COUNT 72 a 139 c 167 g 48 t
ORIGIN

alignment_scores:
Quality: 757.00 Length: 142
Ratio: 5.331 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 98.592

alignment_block:
US-08-962-560A-4 x AA957508/rev ..

Align seg 1/1 to reverse of: AA957508 from: 1 to: 426

```

33 serProAlaAlaProValArgProArgProCysProAlaValProAlaBr 49
|||||
426 TGGCGGCGGCGGCGGCGGCGGCTTCCCGGCGGCGGCGGCGGCGG 377
|||||
49 oAlaProGlyAspThrHisPheArgThrPheArgSerHisSerAspTyrA 66
|||||
376 GGTCTCCGCGGCGGACACTTCCTCCGACCTTCCGCTCCACACTGTGATACC 327
|||||
66 rGArgLleThArgThrSerAlaLeuAspAlaCysGlyPheTyrTrp 82
|||||
326 GCGGCTACAGCGGACCAAGCGCTCTGAGCGGCTGCGGCTTCTACTGG 277
|||||
83 GlyProLeuSerValHisGlyAlaHisGlyAlaArgLeuArgAlaGluProVa 99
|||||

```

```

|||||
276 GGACCCCTGACGCTGATGGGGCGACAGCGCTGCGGAGCCGCT 227
99 1G1YThPheUValArGAspSerArGlnArGAsnCysPheAla 116
226 GGGACCTCTTGTGGCGCGACAGTCGCCAGCGAACTCTTCTCGCGC 177
116 euserValyMerAlaSerGlyProThrSerLeArValAlHisPheGln 132
176 TCAGGTAAGATGGCTTGGGCCCCACAGACATTCGTGACTTCAG 127
133 AlaGlyArGpHeHsLeuAspGlySerArGlnThrPheAspCysLeuPh 149
126 GCCGGCCCTTCACCTGACGCGCGCGGACCTTCGACTGCTCTT 77
149 eGluLeuLengluHsIYrValAlaAlaProArGArGMeLeuGlyAlaP 166
76 CGAGCTGCTGGAGCACTACGTGGCGCGCGCGCATGTGGGGGCC 27
166 roLeuArGlnArGArGValArGPro 174
26 CACTGGCGCAGCGCGCTGGCGCGC 1

seq_name: gb_est38:AM047382

seq_documentation_block:
LOCUS AM047382 422 bp mRNA EST 18-SEP-1999
DEFINITION UI-M-BH1-ame-e-06-0-UI.s1 NIH_BMAP_M.S2 Mus musculus cDNA clone
ACCESSION AM047382
VERSION AM047382.1 GI:5907911
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 422)
Normalisation and subtraction: two approaches to facilitate gene
discovery
JOURNAL genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT On Jun 5, 1998 this sequence version replaced gi:3189644.
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mestr@mail.nih.gov
Oligo-dT track not found. Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA library preparation: M.B.
Soares Lab clone distribution: NIH BMAP cDNA clones will be made
available by the means that is soon to be determined. When NIH
determines the means for distribution of the BMAP cDNA clones, this
record will be updated accordingly when that means is determined.
Seq primer: M13 forward
POLYA-No.
FEATURES
source
1..422
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH1-ame-e-06-0-UI"
/clone_lib="NIH_BMAP_M.S2"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/Note="Vector: pTR3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not 1; Site_2: Eco RI; The
NIH_BMAP_M.S2 library is a subtracted library derived from
NIH_BMAP_M.S1, which in turn is a subtracted library
derived from a mixture of normalized libraries from ten
regions of the mouse brain (cerebellum, brain stems,
```

```

olfactory bulbs, hypothalamus, cortex, amygdala, basal
ganglia, pineal gland, striatum, hippocampus). The driver
used for subtraction consisted of a pool of 5,000 clones
from the NIH_BMAP_M.S1 library and a pool of 2,000 clones
obtained from non-normalized and normalized mouse brain
spinal cord libraries.
TAG_LIB=NIH_BMAP_M.S2
TAG_TISSUE=corpus-striatum
TAG_SEQ=ACGCC"
BASE COUNT 68 a 135 c 169 g 50 t
ORIGIN
alignment_scores:
Quality: 754.00 Length: 140
Ratio: 5.386 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-962-560A-4 x AM047382/rev ..
Align seq 1/1 to reverse of: AM047382 from: 1 to: 422

35 AlaAlaProValArGProArGProCysProAlaValProAlaProAlaPr 51
|||||
420 GCGGCCCCCGTGCCTCCCGGCGCTGCGCGGCTCCCGACGCCAGCCCC 371
51 OGlyAspThrHisPheArGThrPheArGSerHisSerAspTyrArGArG 68
370 TGGCGACACTCATTCCGACACTCCGCTCCCACTCGATACCGGGCGCA 321
68 leThrArGThrSerAlaLeuLeuAspAlaCysGlyPheTyrTrGlyPro 84
320 TCAGCGGACGACGCGCGCTCCTGGACGCGCTGCTTATTTGGGGAGCC 271
85 LeuSerValHisGlyAlaHisGlyArGLeuArGAlaGluProValGlyTh 101
270 CTGACCGCGACGCGGGCGCGACGACGCGCTGCTCCGACGCCGtGGGCAC 221
101 rPheLeuValArGAspSerArGlnArGAsnCysPhePheAlaLeuSerV 118
220 CTCTGTGGTGGCGCACAGTCCCAAGCAAGCACTGCTTCTTCGCGTCAAGC 171
118 allysMeAlaSerGlyProThrSerLeArGValAlHisPheGlnAlaGly 134
170 TGAAGATGGCTTGGCGGCCCGACGAGCATCCGCGTGCACCTTCAGCGCGC 121
135 ArGpHeHsLeuAspGlySerArGlnThrPheAspCysLeuPheGluLe 151
120 CGCTTCCACTGGACGCGGACGCGGAGACCTTCGACTGCTTTCGAGACT 71
151 uLeuGluHsIYrValAlaAlaProArGArGMeLeuGlyAlaProLeuA 168
70 GCTGAGACACTACGTGGGGCGCGCGCGCATGTGGGGGCCCGCGCTGC 21
168 rgGlnArGArGValArGPro 174
20 GCCAGCGCGCGCGTGGCGCGC 1

seq_name: gb_est36:AI873269
1..422
seq_documentation_block:
LOCUS AI873269 464 bp mRNA EST 01-SEP-1999
DEFINITION wF41C08.x1 Soares.NFL_T_GBC.S1 Homo sapiens cDNA clone
IMAGE:2358158 3' similar to TR:015097 015097 JAB, PARTIAL CDS ;
ACCESSION AI873269
VERSION AI873269.1 GI:5547318
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.
```


alignment_scores:

Quality: 722.00 Length: 141
 Ratio: 5.194 Gaps: 0
 Percent Similarity: 98.582 Percent Identity: 97.163

alignment_block:

US-08-962-560A-4 x AA212316 ..

Align seg 1/1 to: AA212316 from: 1 to: 555

```

1 MetValAlaArgAsnGlnValAlaAlaAspAsnAlaIleSerProAlaAla 17
  |||
133 ATGGAGACGACGACACAGGTGGACGCGACAGATGAGATCTCCCGGACG 182
  |||
17 AGUProArgArgArgSerGluProSerSerSerSerSerSerSer 34
  |||
183 AGAGCCCGACGCGGCGTCAGAGCCCTCCTCCTCCTCCTCCTCCTCCT 232
  |||
34 roAlaAlaProValArgProArgProArgProAlaValProAlaProAla 50
  |||
233 CAGGGGCCCCGCTGCTCCCGGCTCCCGGCGGTCGCCAGCCCGGACGC 282
  |||
51 ProGlyAspThrHisPheArgThrPheArgSerHisSerAspTyrArgArg 67
  |||
283 CCTGGCGACTCACTTCCGACCTTCCGCTCCGCTCCGCTCCGCTCCGCT 332
  |||
67 gIleThrArgThrSerAlaLeuLeuAspAlaCysGlyPheTyrTrpTyr 84
  |||
333 CATACACGCGGACGACGCGCTCCGAGCGCTGCGGCTTCTATTTGGGGAC 382
  |||
84 roLeuSerValHisGlyAlaHisGlyAlaArgLeuAlaGluProValGly 100
  |||
383 CCCGAGACGCTGACGAGGCGGCGACAGCGGCTGCTGCCGAGCCGCTGGC 432
  |||
101 ThrPheLeuValAlaArgAspSerArgGlnArgAsnGlySerpPheAlaLeu 117
  |||
433 ACCTTCTTGGTGGCGGACAGTCGCGCAAGGAACTGTTCTTCCGGGCTCAG 482
  |||
117 rValAlaMetAlaSerGlyProThrSerIleArgValHisPheGlnAla 134
  |||
483 CCGTAAGATGGCTTGGGCGCCAGAGACATCCGCTGCTCATTATCCGCC 532
  |||
134 lYArgPheHisLeuAspGlySer 141
  |||
533 GCGCCTTCACTTGACGCGCAGC 555
  |||

```

seq_name: gb_est28:AI505924

seq_documentation_block:

LOCUS AI505924 417 bp mRNA EST 11-MAR-1999
 DEFINITION vk69g08.x1 Knowles Solter mouse 2 cell Mus musculus cDNA clone
 IMAGE:559966 3' similar to TR:035716 035716 SUPPRESSOR OF CYTOKINE
 SIGNALING-1. [1] ; mRNA sequence.

ACCESSION AI505924

VERSION AI505924.1 GI:4403775

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eutheria; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 1 (bases 1 to 417)
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS

TITLE Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
 Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,
 Rafter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
 Waterston, R. and Wilson, R.
 The Mashu-NCI Mouse EST Project 1999
 Unpublished (1999)

JOURNAL On May 18, 1998 this sequence version replaced gi:3136451.
 Contact: Marra M/Mashu-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800
 Fax: 314 286 1810

Email: mouseest@wustl.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 MGI:548758

This clone was previously sequenced on the 5' end only, this new
 data is from the 3' end
 Possible reversed clone: similarity on wrong strand
 Seq primer: Primer name ambiguous
 High quality sequence stop: 347.

FEATURES

source

1..417
 /organism="Mus musculus"
 /strain="B6D2 F1/J"
 /db_xref="taxon:10090"
 /clone="IMAGE:959966"
 /clone_1lb="Knowles Solter mouse 2 cell"
 /tissue="type=embryo"
 /dev_stage="2-cell"
 /lab_host="DH10B"
 /note="Organ: embryo; Vector: pBluescribe (modified);
 Site 1: MluI; Site 2: SalI; cloned unidirectionally from
 mRNA prepared from 13,500 2-cell stage embryos. Primer:
 SalI(drf): 5'-GGTGCAGCCGCGACCGCTTTTCTTTTCTTTT-3'.
 were cloned into the MluI/SalI sites of a modified
 pBluescribe vector using commercial linkers (NEB).
 Average insert size: 1.2 kb."
 BASE COUNT 51 a 162 c 135 g 69 t
 ORIGIN

alignment_scores:
 Quality: 718.00 Length: 139
 Ratio: 5.319 Gaps: 0
 Percent Similarity: 97.122 Percent Identity: 96.403

alignment_block:

US-08-962-560A-4 x AI505924 ..

Align seg 1/1 to: AI505924 from: 1 to: 417

```

42 ProCysProAlaValProAlaProAlaProGlyAspThrHisPheArgTh 58
  |||
1 CCTGCGCGGCGGCTCCAGCCAGCCCGCTGGCGACACTTCCGCGAC 50
  |||
58 rPheArgSerHisSerAspTyrArgArgIleThrArgThrSerAlaLeu 75
  |||
51 CTTCGCTCCCACTCCGATTACGGGGCATCACGGGAGACACGGGCTCC 100
  |||
75 euAspAlaCysGlyPheTyrTrpGlyProLeuSerValHisGlyAlaHis 91
  |||
101 TGGAGCGCTGCGGCTTCTATTGGGAGCCCTGAGGCTGACGCGGCGCAC 150
  |||
92 GluArgLeuArgAlaGluProValGlyThrPheLeuValArgAspSerAr 108
  |||
151 GAGCGGCTGCTGCGGAGCCCGTGGGACCTTCTTGTGCGGACAGTCG 200
  |||
108 gGlnArgAsnCysPhePheAlaLeuSerValHisMetAlaSerGlyProT 125
  |||
201 CCAAGGAACTGCTCTTCCGCGCTCAGGCTGAAGATGGCTTGGGCCCA 250
  |||
125 hSerIleArgValHisPheGlnAlaGlyArgPheHisLeuAspGlySer 141
  |||
251 CGAGCATCCGCGCTGCACTTCCAGCGGCGCGCTTCCACTTGGACGCGAC 300
  |||
142 ArgGluThrPheAspCysLeuPheGluLeuLeuGluHisTyrValAlaAla 158
  |||
301 CCGGAGACTTGTGACTGCTTTTCAAGTGTGAGCACTACGTGGCGGC 350
  |||
158 aProArgArgMetLeuGlyAlaProLeuArgGlnArgArgValArgProL 175
  |||
351 GCGCGCCGATGAGGGGGGCGCGCTGCGCAGCGCGCGCGGCGCGC 400
  |||

```

175 euglinc1eucysarg 180
 |||||:|||||||
 401 TCGAGAGCTGTGTGCG 417

seq_name: gb_est38:AM048541

seq_documentation_block:

LOCUS AM048541 389 bp mRNA EST 18-SEP-1999
 DEFINITION UI-M-BH1-alk-c-08-0-UI-s1 NIH_BMAP_M_S2 Mus musculus cDNA clone
 UI-M-BH1-alk-c-08-0-UI 3', mRNA sequence.

ACCESSION AM048541
 VERSION AM048541.1 GI:5909070

KEYWORDS EST.

SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 1 (bases 1 to 389) Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477

COMMENT On Mar 9, 1998 this sequence version replaced g1:2946940.

Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890

Email: mestr@mail.nih.gov
 Oligo-dt track not found, Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.B.
 Soares Lab Clone distribution: NIH BMAP cDNA clones will be made
 available by the means that is soon to be determined. When NIH
 determines the means for distribution of the BMAP cDNA clones, this
 record will be updated accordingly when that means is determined.

Seq primer: M13 Forward

POLYA-NO.

FEATURES Location/Qualifiers

source

1..389 /organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UI-M-BH1-alk-c-08-0-UI"

/clone.lib="NIH_BMAP_M_S2"

/dev_stage="27-32 days"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pRTT3D-Pac (Pharmacia) with a modified

polylinker. Site_1: Not I; Site_2: Eco RI; The

NIH_BMAP_M_S2 library is a subtracted library derived from

NIH_BMAP_M_S1, which in turn is a subtracted library

derived from a mixture of normalized libraries from ten

regions of the mouse brain (cerebellum, brain stems,

olfactory bulbs, hypothalamus, cortex, amygdala, basal

ganglia, pineal gland, striatum, hippocampus). The driver

used for subtraction consisted of 5,000 clones

from the NIH_BMAP_M_S1 library and a pool of 2,000 clones

obtained from non-normalized and normalized mouse brain

spinal cord libraries.

TAG_LIB=NIH_BMAP_M_S2

TAG_TISSUE=corpus striatum

TAG_SEQ=ACGCGC

BASE COUNT 65 a 123 c 151 g 50 t

ORIGIN

alignment_scores:

Quality: 687.00 Length: 129

Ratio: 5.326 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 99.225

alignment_block:

US-08-962-560A-4 x AM048541/rev ..

Align seg 1/1 to reverse of: AM048541 from: 1 to: 389

46 ValProAlaProAlaProGlyAspThrHisPheArgThrPheArgSerHis 62
 |||||:|||||
 387 GTCCACACCCCGCCGCTGGCGACACTCTCCGACCTCCGCTCCCA 338
 62 sSerAspTyrArgArgGlyIleThrArgThrSerAlaLeuAlaPalaCysG 79
 |||||:|||||
 337 CTCGCATTACCGCGCATCACCGGACGCGCGCTCTGACGCGCTCG 288
 79 LpPheTyrTrpGlyProLeuSerValHisGlyAlaHisGlyArgLeuArg 95
 |||||:|||||
 287 GCTTCTATTGGGAGCCCTGACGCTGCACGGGGCCACGAGGGCTGCT 238
 96 AlaGluProValGlyThrPheLeuValArgAspSerArgGlnArgAsnCy 112
 |||||:|||||
 237 GCCGACCCCGCTGGCGACCTTCTGTGGCGGACAGTCCCAACGAACTG 188
 112 spPheAlaLeuSerValLysMetAlaSerGlyProThrSerIleArgV 129
 |||||:|||||
 187 CTTCTTCGCGCTCACGCTGAAGATGCTTCGGGCCCGACGACATCCCG 138
 129 AlHisPheGlnAlaGlyArgPheHisLeuAspGlySerArgGlyThrPhe 145
 |||||:|||||
 137 TCCACTCCAGCGCGCGCTTCCTCCACTTGCAGCGGACCGCGAGACTTC 88
 146 AspCysLeuPheGluLeuLeuGluHisTyrValAlaAlaProArgArg 162
 |||||:|||||
 87 GACTCGCTTTCGAGCTGCTGGAGCAGTACGAGCGCGCGCGCGCAT 38
 162 tLeuGlyAlaProLeuArgGlnArgArgValArgPro 174
 |||||:|||||
 37 GTTGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1

seq_name: gb_est28:A1502911

seq_documentation_block:

LOCUS A1502911 393 bp mRNA EST 11-MAR-1999
 DEFINITION UI-R-C1-Kn-f-07-0-UI-s1 UI-R-C1 Rattus norvegicus cDNA clone
 UI-R-C1-Kn-f-07-0-UI 3', mRNA sequence.

ACCESSION A1502911

VERSION A1502911.1 GI:4400762

KEYWORDS EST.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 1 (bases 1 to 393) Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

COMMENT On Mar 10, 1998 this sequence version replaced g1:2948661.

Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: msoares@blue.ueg.uiowa.edu

Oligo-dt track not found, Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.B.
 Soares Lab Clone distribution: clones will be available through

Research Genetics (www.resgen.com)

Seq primer: M13 Forward.

FEATURES Location/Qualifiers

source

1..393

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

```

/bd_xref="taxon:10116"
/clone="UI-R-C1-ko-f-07-0-UI"
/clone.lib="UI-R-C1"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/Note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI. The UI-R-C1
library is a subtracted library derived from the UI-R-C0
library, which is a subtracted library derived from the
UI-R-A1 and UI-R-R1 libraries. The UI-R-A1 library
consisted of a mixture of individually tagged normalized
libraries constructed from rat placenta, adult lung,
brain, liver, kidney, heart, spleen, ovary, and muscle.
The UI-R-E1 library consisted of a mixture of
individually tagged normalized libraries constructed from
8, 12 and 18-day embryo. The tag is a string of 3-5
nucleotides present between the Not I site and the
oligo-dr track which allows identification of the library
of origin of a clone within the mixture. The subtracted
library (UI-R-C1) was constructed as follows: PCR
amplified cDNA inserts from UI-R-C0 clones from which 3'
ESTs had been derived was used as a driver in a
hybridization with the UI-R-C0 library in the form of
single-stranded circles. The remaining single-stranded
circles (subtracted library) was purified by
hydroxyapatite column chromatography, converted to
double-stranded circles and electroporated into DH10B
bacteria (Life Technologies) to generate the UI-R-C1
library. This procedure has been previously described
(Bonaldi, Lennon and Soares, Genome Research 6: 791-806,
1996)."

```

```

BASE COUNT      66 a      128 c      149 g      50 t
ORIGIN
alignment_scores:
  Quality: 680.00      Length: 127
  Ratio: 5.354      Gaps: 0
  Percent Similarity: 100.000      Percent Identity: 100.000

```

alignment_block:

US-08-962-560A-4 x A1502911/rev ..

Align seg 1/1 to reverse of: A1502911 from: 1 to: 393

```

48 AAlaProAlaProGlyAspThrHisPheArgThrPheArgSerHisSerAs 64
|||||
381 GCCCGGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 332
64 PTYRARGTGLIETHRARGTHSERALALALALALALALALALALALALALAL 81
|||||
331 TTACCGGCGCATCGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTTCT 282
81 YTRPGIYPRoleuservAlHisGlyAlAHISGLIARGLIARGLIARGLIAR 97
281 ACTGGGACCCCTGAGCGTGCATGCGGCGGCGGCGGCGGCGGCGGCGGAG 232
98 ProValGlyThrPheLeuValArgAspSerArgGlnArgAspCysPhePh 114
|||||
231 CCCGTGGGCGACTTCTTGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCTCT 182
114 GAAlaLeuSerValIlyMetAlaSerGlyProThrSerIleArgValHisP 131
181 CGCGCTACGCGTGAAGATGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCT 132
131 heGlnAlaIlyArgPheHisLeuAspGlySerArgGlnThrPheAspGly 147
131 TCCAGGCGCGCGCTTCCACCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCT 82
148 LeuPheGlnLeuLeuGlnHisTyrValAlaAlaProArgMetLeuGln 164
81 CTCTTCGAGCTGCTGAGCAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCTGG 32

```

```

164 YAlaProLeuArgGlnArgValArgPro 174
|||||
31 GCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1

```

seq_name: gb_est16:AA547459

```

seq_documentation_block:
LOCUS      AA547459      475 bp      mRNA
DEFINITION VK65908.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone
IMAGE:59966 5', mRNA sequence.
ACCESSION  AA547459
VERSION    AA547459.1 GI:2308750
KEYWORDS  EST.
SOURCE     house mouse.
ORGANISM  Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 475)
Matta,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisler,S., Kucada,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMNI Mouse EST Project
Unpublished (1996)
On Apr 14, 1993 this sequence version replaced gi:692678.
Contact: Matta M/Mouse EST Project
WashU-HMNI Mouse EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LIND; contact the
IMAGE Consortium (info@image.lind.gov) for further information.
MGI:548758

```

TITLE
JOURNAL
COMMENT

FEATURES
source

```

location/Qualifiers
1..475
/organism="Mus musculus"
/strain="B6D2 F1/7"
/bd_xref="taxon:10090"
/clone="IMAGE:959966"
/clone.lib="Knowles Solter mouse 2 cell"
/tissue_type="embryo"
/dev_stage="2-cell"
/lab_host="DH10B"
/Note="Organ: embryo; Vector: pBluescribe (modified);
site 1: MluI; site 2: SalI; cloned unidirectionally from
mRNA prepared from 13,500 2-cell stage embryos. Primer:
SalI(dT): 5'-CGGTGACCGTGCAGCCGTTTCTTTTCTTTT-3'.
were cloned into the MluI/SalI sites of a modified
pBluescribe vector using commercial linkers (NEB).
Average insert size: 1.2 kb."

```

```

BASE COUNT      64 a      171 c      148 g      92 t
ORIGIN

```

```

alignment_scores:
  Quality: 607.00      Length: 159
  Ratio: 4.245      Gaps: 3
  Percent Similarity: 89.937      Percent Identity: 88.050

```

alignment_block:

US-08-962-560A-4 x AA547459 ..

Align seg 1/1 to: AA547459 from: 1 to: 475

```

54 ThrHisPheArgThrPheArgSerHisSerAspTyrArgArgIleThrArg 70
|||||
2 ACTCACTTCGCGACCTTCGCTCCGCTCCGATTCGATTCGCGGCGGCGGCGG 51
70 gThrSerAlaLeuLeuAspAlaGlyGlyPheTyrTTrpIlyProLeuServ 87

```

```
|||||
52 GACCGACCGCGCTCTGAGCGCTGCGG..TTCTATTGGGAGACCTTG..AGCG 99
87 aIHISGIVAlaHISGIVAlaGLeuArgAlaGluProValGlyThrPheLeu 103
100 TGCACGGGGCGCGACGAGCGGTGCTGCCGAGCCGCTGGGACCTTCTTG 149
104 ValArgAspSerArgGlnArgAsnGlyPheAlaLeuSerValLysme 120
150 TCGCGACAGCTCG...CAACGGAAGTCTCTTCGCGCTCAC..GTGAGAT 195
120 ValAspGlyProThrSerIleArgValHisPheGlnAlaGlyArgPhe 137
196 GGCCTTGGGCGCCAGAGCATCCGGCTGCATCCAGGCGCGCGCTTCC 245
137 ILeuAspGlySerArgGluThrPheAspCysLeuPheGluLeuGlu 153
246 ACTTGAGACGCGACCGCGAGACCTTCGACTGCTTTCGAGCTGCTGAG 295
154 HistyrValAlaAlaProArgArgMetLeuGlyAlaProLeuArgGln 170
296 CACTACGTGGCGCG...CGCGGATGTGGGGCGCGCTGCGCAGCGCGCG 342
170 GArgValArgProLeuGlnIleuGlyArgGlnArgIleValAlaAla 187
343 TGC.....GCGCTGCAGAGAGCTGTGTGCCAGCGCATGTGGCGCG 385
187 aIGlyArgGluAsnLeuAlaArgIleProLeuAsnProValLeuArg 203
386 TGGGTTCGGAGAACCTGGCGCGCATCCCTTAACCGGTAATCGTGAC 435
204 TyrLeuSerSerPheProPheGlnIle 212
436 TACCTGAGTTCTTCCCTTCAGATC 462
```